Curious inferences: reply to Sun & Firestone on the Dark Room Problem

Article (Accepted Version)

Curious Inferences: Reply to Sun & Firestone on the Dark Room Problem

Anil K. Seth[1,2,4*], Beren Millidge[3], Christopher L. Buckley[1], and Alexander Tschantz[1,2]

[1] Department of Informatics, University of Sussex, Brighton BN1 9QJ, UK
[2] Sackler Centre for Consciousness Science, University of Sussex, Brighton, BN1 9QJ, UK
[3] Canadian Institute for Advanced Research (CIFAR), Azrieli Program on Brain, Mind, and Consciousness
[4] Department of Informatics, University of Edinburgh, EH8 9AB, UK

*correspondence: a.k.seth@sussex.ac.uk

Keywords: expected free energy; predictive processing; active inference; the free energy principle;

Sun & Firestone [1] present a challenge to predictive processing (PP) accounts of brain function by reviving the Dark Room Problem – the idea that if agents are mandated to minimise prediction error, the best thing for them to do is to seek out highly predictable environments where nothing changes, and stay there. They argue that standard responses to this challenge have the potential to render the PP account untestable and explanatorily empty. We disagree.

One standard response is that Dark Room type environments are intrinsically surprising, given the homeostatic imperatives of living organisms. One might worry that this response solves nothing, since it merely redefines what counts as 'surprising' for an agent. The reply by Van de Cruys and colleagues relieves us of this worry by highlighting the principled role of 'optimistic predictions' in driving actions [2].

A second, and related, response is that increasing prediction error in the short term – for example by leaving a Dark Room to engage in curious exploration – may help to reduce prediction error in the long run. Sun & Firestone argue that this response is also inadequate because "not all motivations that drive us from dark rooms reduce to instrumentally valuable exploration, even over the long-term" (p. 346). To support their point, they note that some distinctive – though rare – human behaviours, such as riding rollercoasters and reading poetry, do not seem to deliver instrumental (goal-oriented) benefit, even over the long-term.

Fortunately, this objection loses its force when the role of *action* is properly taken into consideration, as exemplified by 'active inference' formulations of PP [3, 4].

Action holds a special position in active inference. Unlike inferences about sensory states and internal states, (proprioceptive) inferences leading to actions can directly change the environment. From the perspective of the agent, sequences of actions thus 'change the future'. Given that an agent is compelled to minimise long-term prediction error, it is therefore also mandated to reduce its uncertainty about the world, so that it can better minimise the discrepancy between expected and actual sensory data across temporally-extended sequences of actions. Such an agent will therefore engage in epistemic, information-seeking actions – such as leaving a Dark Room – even though such actions may transiently increase short-term prediction error. In short, in order to minimise surprise *in the future*, an agent needs to be a curious, sensation-seeking agent *in the present*.

Why does this response not fall foul of Sun & Firestone's critique that it simply redefines what is 'surprising'? The reason is that minimisation of long-term prediction error can be formalised in a way that makes specific, testable predictions. The formalism that makes this possible is the 'free energy principle' (FEP) which generalises PP to propose that organisms minimise the 'free-energy', a tractable (i.e., measurable from the perspective of the agent) upper-bound on the long-term average of sensory entropy, which generalises the notion of prediction error.

When it comes to Dark Rooms, agents must minimise free-energy over long temporal horizons – over the long term. Mathematically, this means that agents must minimize not free energy *per se*, but the 'expected free energy' – a quantity which can be formalized in various ways (see Box 1). Minimizing expected free energy entails minimizing a (negative) expected information gain term, which rewards sampling those novel environmental states which (are predicted to) induce a large divergence between prior and posterior beliefs. This is why long-term free-energy-minimizing agents are intrinsically drawn towards novel experiences (and thus out of Dark Rooms) that reduce uncertainty about the world, even at the expense of temporarily higher prediction error. Importantly, the free-energy functional (a function of a function) intrinsically balances this trade-off between immediate and long-term free-energy minimization.

Formalising the situation in this way leads to testable hypotheses. For example, by performing variational inference in computational models that have parameters corresponding to beliefs about actions, one can make specific predictions about epistemic actions such as eye-movements [5]. By incorporating learning, one can also make predictions about the biases that may accrue to an agent's beliefs about the world, as it attempts to minimise expected free energy [6]. By reconstruing goals and rewards as prior expectations, these models can also make fine-grained predictions about the dynamics of reinforcement learning [7].

Will this approach extend to explain rollercoaster-riding and poetry-reading? In the details, perhaps not. But this is not a failure, nor is it – as Sun & Firestone suggest – a concern over the explanatory reach of PP and the FEP. Here, it is important to recognise that the FEP is a framework, not a testable hypothesis in and of itself. The lasting value of the FEP will lie in the fecundity with which it generates virtuous circles of testable hypotheses – such as those deriving from models of expected free energy – and not with any specific attempt to 'prove' or 'falsify' it [8].

[800 words]

Box 1: Extending free energy into the future

We have highlighted that when agents minimise free energy across long temporal horizons, this naturally induces information-seeking (Dark Room escaping) behaviour. What does 'minimising free energy into the future' mean? The idea is to conceptualise inference as operating over sequences of observations, states, and actions extending into the future. This leads to an imperative to minimise 'expected free energy'(EFE), which quantifies the total free energy of a sequence of observations and actions.

Mathematically, the EFE can be expressed in several ways (through different free-energy functionals) [9], each of which can be decomposed into 'instrumental' terms, which promote the immediate fulfilment of prior beliefs, and – crucially – 'epistemic' terms, which promote exploration of novel environmental contingencies. Importantly these epistemic terms arise naturally out of the mathematical formalism, instead of being bolted on 'ad-hoc', and they only arise when performing inference over temporally extended sequences. This is because the epistemic terms function as Bayes-optimal mediators of the trade-off between short and long-term free-energy minimization. Moreover, the properties of different free-energy functionals can be distinguished and tested empirically, thus leading to testable predictions about the specific functionals that agents employ when they escape from Dark Rooms.

[180 words]

References

1. Sun, Z. and C. Firestone, *The Dark Room Problem.* Trends Cogn Sci, 2020. **24**(5): p. 346-348.
2. Van der Cruys, S., K.J. Friston, and A. Clark, *Controlled optimism: Reply to Sun and Firestone on the Dark Room Problem.* Trends Cogn Sci, 2020.
3. Friston, K.J., et al., *Action and behavior: a free-energy formulation.* Biological Cybernetics, 2010. **102**(3): p. 227-60.

4. Buckley, C. L. , et al., *The free energy principle for action and perception: A mathematical review.* Journal of Mathematical Psychology, 2017. **81**: p. 55-79.
5. Mirza, M.B., et al., *Human visual exploration reduces uncertainty about the sensed world.* PLoS One, 2018. **13**(1): p. e0190429.
6. Tschantz, A., A.K. Seth, and C. L. Buckley, *Learning action-oriented models.* PLoS Comput Biol, 2020.
7. Tschantz, A., et al., *Reinforcement learning through active inference.* 2020.
8. Lakatos, I., *The methodology of scientific research programmes: Philosophical papers*. 1978, Cambridge: Cambridge University Press.
9. Milidge, B., A. Tschantz, and C. L. Buckley, *Whence the expected free energy?* 2020.